"EXPRESS MAILING"

Mailing Label No.: EV 393967769 US
Date of Deposit: January 28, 2004

METHOD AND CIRCUIT FOR REDUCING QUANTIZER INPUT/OUTPUT SWING IN A SIGMA-DELTA MODULATOR

TECHNICAL FIELD

[0001] The present invention relates to sigma-delta, sometimes referred to as delta-sigma, modulators used in analog-to-digital converter circuits, specifically a circuit and method for reducing the quantizer input/output swing in a sigma-delta modulator which enables the modulator to have a larger input/output signal swing before the signal at the input/ouput of the quantizer and the output of the last integrator saturates.

BACKGROUND OF THE INVENTION

[0002] Most signals of practical interest, such as speech, radar signals, sonar signals and various communication signals such as audio and video signals are analog. Often it is desirable to process these analog signals by digital means. To process an analog signal by digital means, it is necessary to convert the analog signal into digital form.

[0003] In analog-to-digital conversion, a smooth and continuous analog signal is converted or quantized into a sequence of numbers having finite precision with discrete values at discrete times. The devices that perform this task are called analog-to-digital ("A/D") converters ("ADC"). Modern demands for precise, fast processing of such digital signals coupled with the advances in very-large-scale integrated circuit ("VLSI") technology have resulted in the necessity for cost-effective and highly accurate ADC. There are generally two types of ADC commonly used for high-resolution application, sigma-delta modulator ADC and successive approximation

register ("SAR") ADC. Sigma-delta modulator ADC have certain advantages over SAR ADC, when used for oversampled noise-shaping, including inherent linearity and high tolerance to circuit imperfection.

[0004] A sigma-delta ADC is generally comprised of (1) a sigma-delta modulator followed by (2) a digital decimation filter and a (3) low-pass filter. The sigma-delta modulator consists of an analog filter/integrator and an n-bit quantizer enclosed in a feedback loop via an m-bit DAC. The feedback DAC is implemented as an array of finite capacitors configured such that a selected number of these capacitors release their electrical charge into a summing junction that produces an equivalent analog output signal of the m-bit digital input code. The quantizer consists of an array of finite comparators in parallel such that each comparator compares the analog signal at the quantizer input to a reference voltage associated to its own output level. The reference voltage for each comparator in the array is equally spaced by the number of comparators in the array within the A/D positive full-scale and negative full-scale. A common clock latch/triggers the output of the comparators, such that each comparator generates a logic "high" (1) or a logic "low" (0) level, with the parallel output of the comparator representing a digital "thermometer code" equivalent to the analog signal at the quantizer input. This "thermometer code" is digitally processed to generate an n-bit digital word representing the converted analog signal that is the output of the modulator and also fed back into the analog filter loop via the feedback DAC. The error between the modulator analog input signal and the modulator digital output code due to the quantizer is the quantization noise

[0005] Together with the analog filter, the feedback loop acts to attenuate the quantizing noise at low frequencies while emphasizing the high frequency noise. Since the signal is sampled at a frequency much higher than the Nyquist rate, i.e., a higher oversampling

ratio ("OSR"), high frequency quantization noise can be removed without affecting the signal band of interest by means of a digital low-pass filter operating on the output of the sigma-delta modulator. The digital low-pass filter is implemented together with the decimation filter to convert the data from a high rate, low resolution bit stream to a lower rate, higher resolution digital output through oversampling and decimation. Hence, in operation, the sigma-delta modulator oversamples the signal and shapes the noise out of the band of interest.

[0006] In the past, the single-bit (1-bit quantizer with 1-bit feedback DAC) sigma-delta modulator has been the preferred architecture due to its low sensitivity to analog component matching by having a large OSR. This architecture has dominated the market for very high resolution but lower speed ADCs. Recently, there has been a desire for high speed, high resolution sigma-delta ADCs. Disadvantageously, sampling frequency is limited by the capability of the device and silicon process technology. Therefore, in order to achieve a higher output data rate, the ADC must have a lower OSR. Most of the higher speed sigma-delta modulators are taking advantage of the multi-bit and multi-stage noise shaping ("MASH") schemes in order to compensate for the degradation of the signal-to-noise ratio ("SNR") due to lower OSR constraints. In the multi-bit scheme, the n-bit and m-bit of the quantizer and feedback DAC respectively is more than 1. In the multi-stage scheme, more than one modulator stage is cascaded.

[0007] One of the architectural limitations for SNR and spurious-free-dynamic range ("SFDR") degradation with increasing input signal swing is due to the maximum signal swing limitation. Any given circuit block has its own maximum input signal swing limitation. An input signal that exceeds this limit is typically clipped or saturated. Signal clipping and saturation causes both harmonic and non-harmonic tones. For example, the maximum signal swing of the

switched-capacitor ("SC") switch is limited by the gate driving voltage, and the maximum signal swing of the integrator is limited by the operational transconductance amplifier ("OTA") output swing which is usually limited by the power supply. The maximum signal swing limitation in a quantizer is due to the finite quantization level or the 0dB full-scale range of the A/D converter. As previously noted, the n-bit of the quantizer and the m-bit of the feedback DAC have finite quantization levels

SUMMARY OF THE INVENTION

[0008] The present invention achieves technical advantages as a circuit arrangement and method by which the input/output swing of the quantizer in the sigma-delta modulator is reduced or kept within the limit of the finite quantization levels. Advantageously, the quantizer output swing reduction circuit and method of the present invention enables the modulator to have a larger input/output swing range without degrading the SNR and SFDR performance due to quantizer saturation and clipping.

[0009] A gradually increasing signal introduced into an A/D converter using the circuit and method of the present invention will not clip or saturate at the output of the quantizer as compared to a conventional sigma-delta modulator. Therefore, the circuit and method avoids large signal clipping that causes harmonic distortion and tones which increase the noise floor in the in-band frequency spectrum.

[0010] These and many other advantages related to the improvements of the sigma-delta modulator will become apparent to persons skilled in the relevant arts through careful reading of the disclosure and claims presented herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] For a better understanding of the invention including its features, advantages and specific embodiments, reference is made to the following detailed description along with accompanying drawings in which:

[0012] Figure 1(a) is a block diagram of a first conventional second order sigma-delta modulator;

[0013] Figure 1(b) illustrates a quantizer error/noise model;

[0014] Figure 2 is a block diagram of a first embodiment of the present invention, using an arbitrary signal, Vx, as implemented with a second order sigma-delta modulator;

[0015] Figure 3 is a block diagram of a second embodiment of the present invention, using the modulator output signal, Yout, as Vx, as implemented with a second order sigma-delta modulator;

[0016] Figure 4(a) is a block diagram of an embodiment of the present invention, using the modulator output signal, Yout, as Vx, as implemented with a second order sigma-delta modulator, including a non-delay first integrator and delay second integrator, wherein all the gain co-efficients a=b=c=d=1;

[0017] Figure 4(b) is a block diagram of a conventional second order sigma-delta modulator with non-delay first integrator and delay second integrator, wherein all the gain co-efficients a=b=c=d=1;

[0018] Figures 5(a)-(d) are plots of the probablity of occurrences at each voltage values of a sine wave input over a period of time at different output stages;

[0019] Figure 6 is a composite plot of the in-band frequency spectrum of the output of a second order sigma-delta modulator using the present invention seen in Figure 4(a) compared to a second order sigma-delta modulator without the present invention as seen in Figure 4(b);

[0020] Figures 7(a) and (b) are plots of the SNR and SFDR performance of a second order sigma-delta modulator using the present invention as seen in Figure 4(a) compared to a second order sigma-delta modulator without the present invention as seen in Figure 4(b) with respect to modulator input signal swing;

[0021] Figure 8(a) is a block diagram of an embodiment of the present invention, using the modulator output signal, Yout as Vx, as used in a second conventional second order sigma-delta modulator with delay in both first and second integrator and the gain co-efficients, a=c=1/2 and b=d=2;

[0022] Figure 8(b) is a block diagram of a conventional second order sigma-delta modulator with delay in both first and second integrator and the gain co-efficients, a=c=1/2 and b=d=2;

[0023] Figure 9(a) is a block diagram of an embodiment of the present invention, using the modulator output signal, Yout as Vx, as implemented in a first order sigma-delta modulator with a delay in the first integrator and the gain co-efficients, a=b=1;

[0024] Figure 9(b) is a schematic diagram of a conventional first order sigma-delta modulator with a delay in the first integrator and the gain co-efficients, a=b=1;

[0025] References in the detailed description correspond to like references in the figures unless otherwise noted.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

[0026] While the making and using of various embodiments of the present invention are discussed in detail below, it should be appreciated that the present invention provides many applicable inventive concepts which can be embodied in a wide variety of specific contexts.

[0027] The present invention achieves technical advantages as a circuit arrangement and method by which the input/output swing of the quantizer in a sigma-delta modulator is reduced compared to a conventional sigma-delta modulator with the same input signal amplitude. The present invention improves the performance of an ADC by taking into account the quantizer output swing. With respect to the feedback DAC finite quantization limitation, it is assumed that there is no feedback DAC saturation limitation. Any input to the quantizer that is larger than the full-scale range or maximum quantization level will clip or saturate to the digital full-scale code. Signal clipping introduces harmonic distortion and tones that raise the noise floor. Therefore, the SNR and SFDR of sigma-delta modulators, even the ideal ones without circuit imperfection, start to degrade as the input signal swing increases near the 0dB full-scale range. This is because the analog signal at the input of the quantizer consists of not only the delayed input signal component but also another component which is a function of the quantization error/noise. The present invention enables the ADC to operate with an input signal near full-scale and beyond the 0dB full-scale range of the ADC without degradation in performance, assuming no saturation limitation on the feedback DAC or any other circuit block

limitations. One embodiment of the present invention disclosed herein is implemented using a second order sigma-delta modulator. The second order sigma-delta modulator has inherent stability and a higher order of noise shapping, thus increasing the SNR of the ADC.

[0028] Referring now to the Figures, Figure 1(a) is a conventional second order sigma-delta modulator 100 shown in block diagram, which consists of the forward-path gain "a" 118 and gain "b" 121, feedback-path gain "c" 119 and gain "d" 120, first feedback DAC 124, second feedback DAC 125, first integrator 117 with transfer function H1 and second integrator 103 with transfer function H2 followed by a quantizer 105. As seen in Figure 1(b), the quantizer 105 can be modeled as an addition of the quantization error or noise εq, 122 to the input Y2 104 of the quantizer 105. For any second order sigma-delta modulator, the condition for ab=1, bc=1, and c=a holds.

[0029] A conventional sigma-delta modulator 100 used in an ADC has architectural limitations that result in SNR and SFDR degradation when input signals are increased near the 0dB full-scale voltage. This is primarly due to the limitation of the finite quantization levels. Referring to Figure 1(a) and Figure 1(b), the first integrator 117 output Y1 128, second integrator 103 output Y2 104, the quantizer 105 output Y3 106, and the modulator 100 output Yout 106, can be expressed as equations [1], [2], [3] and [4] below:

$$Y1 = Vin \cdot \left(\frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2}\right) \cdot \left(\frac{1}{H2} + d\right) - \varepsilon_q \cdot \left(\frac{H1}{1 + H1 \cdot H2 + d \cdot H2}\right) \qquad [1]$$

$$Y2 = Vin \cdot \frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2} - \varepsilon_q \cdot \left(\frac{H1 \cdot H2 + d \cdot H2}{1 + H1 \cdot H2 + d \cdot H2}\right) \qquad [2]$$

$$Y3 = Yout = Vin \cdot \frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2} + \varepsilon_q \cdot \left(\frac{1}{1 + H1 \cdot H2 + d \cdot H2}\right) \qquad [3]$$

$$Yout = Vin \cdot \frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2} + \varepsilon_q \cdot \left(\frac{1}{1 + H1 \cdot H2 + d \cdot H2} \right) \quad [4]$$

[0030] It can be seen that if a signal, Vin, with full-scale swing is applied at the modulator 100 input 101, the swing at the quantizer 105 output Y3 106 will swing more than the full-scale input signal, Vin, due to the addition of the quantization noise term. The conventional modulator 100, having lower quantization bits, will clip or saturate sooner than the one with higher quantization bits. However, no matter how high the quantization level, unless it is infinite, the signal through quantizer 105 will saturate before the input/output signal of the modulator 100 reaches 0dB full-scale, causing the increase in both the distortion and the noise floor.

[0031] As seen in Figure 2, there is disclosed an embodiment of a sigma-delta modulation circuit 200 in accordance with the present invention. It comprises the conventional sigma-delta modulation circuit 100 of Figure 1(a) and a quantizer swing reduction block 210. The quantizer swing reduction block 210 comprises an input signal Vx 216, a signal processing block 214 with transfer function H3 and another signal processing block 215 with transfer function H2*H3. The signal before the input 202 of the second integrator 203 with transfer function H2 is subtracted by signal 211, which is Vx 216 passed through transfer function H3 214. This is done such that the swing at the output Y2 208 of the second integrator 203 and similarly at the input 208 of the quantizer 205 will be reduced by the product of H2, H3, and Vx. Another signal 212 which is Vx 216 passed through the product of transfer function H2, and H3, 215 is added to the output 209 of the quantizer 205, before the modulator 200 output Yout 206. This is done such that the swing at the output 206 of the modulator 200 is compensated and remains the same as the output 106 of the modulator 100 of Figure 1, without

the quantizer swing reduction circuit 210. Further referring to Figure 2, the first integrator 217 output Y1 228, second integrator 203 output Y2 208, the quantizer 205 output Y3 209, and the modulator 200 output Yout 206, can be expressed as equations [5], [6], [7] and [8] below:

$$Y1 = Vin \cdot \left(\frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2}\right) \cdot \left(\frac{1}{H2} + d\right) - \varepsilon_q \cdot \left(\frac{H1}{1 + H1 \cdot H2 + d \cdot H2}\right) \quad [5]$$

$$Y2 = Vin \cdot \frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2} - \varepsilon_q \cdot \left(\frac{H1 \cdot H2 + d \cdot H2}{1 + H1 \cdot H2 + d \cdot H2}\right) - Vx \cdot H2 \cdot H3 \quad [6]$$

$$Y3 = Vin \cdot \frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2} + \varepsilon_q \cdot \left(\frac{1}{1 + H1 \cdot H2 + d \cdot H2}\right) - Vx \cdot H2 \cdot H3 \quad [7]$$

$$Yout = Vin \cdot \frac{H1 \cdot H2}{1 + H1 \cdot H2 + d \cdot H2} + \varepsilon_q \cdot \left(\frac{1}{1 + H1 \cdot H2 + d \cdot H2}\right) \quad [8]$$

[0032] The quantizer output swing reduction circuit 210 of the present invention advantageously enables the modulator 200 to have a larger input/output signal swing range without degrading the SNR and SFDR performance. Using circuit 210 of the present invention, the signal is not clipped or saturated by the quantizer 205 as early as the conventional sigma-delta modulator 100. The quantization output swing reduction method and circuit 210 can be used with any order sigma-delta modulator. The choice of transfer function, H3, depends on how much reduction is desired. In other words, H3 is not necessarily limited by a particular transfer function, it can be any arbitrary transfer function that can achieve any amount of signal swing reduction desired at the quantizer 205 output Y3 209.

[0033] Assuming that signal Vx 216, is equal to input signal Vin 201, the transfer function H3 214 can be chosen such that the product of H2 and H3 and the input signal Vin 201, can be used to completely eliminate the term due to the input signal Vin 201 at the output Y3 209 of the quantizer 205.

$$H3 = \frac{H1}{1 + H1 \cdot H2 + d \cdot H2} \qquad [9]$$

[0034] Referring now to Figure 3, modulator 300 output Yout 306, can also be used as Vx because Yout 306 is a close approximation of Vin 301 and Yout 306 is readily available. In this case, although it is not totally eliminated, the term in equation [3] related to the input voltage Vin 301, is reduced. Because Yout 306 is available in the digital domain, the whole quantizer output swing reduction circuit block 310 can be implemented in the digital domain 329. The swing reduction signal 311 can be combined 323 digitally with the output signal 306 before negatively fedback into the input 302 of the analog integrator/filter via the second feedback DAC 325.

[0035] Figure 4(a) illustrates using Yout 406 as Vx to implement the quantizer swing reduction method and circuit of the present invention with a conventional second-order sigma-delta modulator which comprises first integrator/filter 417 with transfer function $H1 = \frac{1}{1 + Z^{-1}}$, a second integrator/filter 403 with transfer function $H2 = \frac{Z^{-1}}{1 + Z^{-1}}$ and gain co-efficients, "a" 418, "b" 421, "c" 419 and "d" 420 set to unity. The transfer function in the quantizer swing reduction block 410 is chosen such that $H3 = 1 - Z^{-1}$ 414 and $H2 \cdot H3 = Z^{-1}$ in order to obtain the desired node voltage at the first integrator 417 output Y1 428, second integrator 403, output Y2 408, quantizer 405 output Y3 409, and the modulator 400 output Yout 406 as expressed in equations [10], [11], [12] and [13] below:

$$Y1 = Vin - \varepsilon_q \cdot \left(1 - Z^{-1}\right) \qquad [10]$$

$$Y2 = Vin \cdot Z^{-1} \cdot \left(1 - Z^{-1}\right) + \varepsilon_q \cdot \left[\left(1 - Z^{-1}\right)^3 - 1\right] \qquad [11]$$

$$Y3 = Vin \cdot Z^{-1} \cdot \left(1 - Z^{-1}\right) + \varepsilon_q \cdot \left(1 - Z^{-1}\right)^3 \quad [12]$$

$$Yout = Vin \cdot Z^{-1} + \varepsilon_q \cdot \left(1 - Z^{-1}\right)^2 \quad [13]$$

Referring to equation [12], the signal reduction at the quantizer 405 output Y3 409, is achieved by means of first order shaping on the original quantizer output swing (i.e. both the delayed modulator 400 input signal, Vin, term and the second order quantization noise, εq, shaping term).

[0036] As seen in Figures 4(a) and 4(b), a conventional second order sigma-delta modulator with 400 and without 430 the quantizer output swing reduction scheme 410 are simulated with the following parameters: OSR = 32, sinusoidal input signal with amplitude, Vin = 0dB (full-scale), and frequency, Fin = 38.7573242 kHz, sampling frequency, FS = 40 MHz, quantizer resolution, N = 3-bit unless specified on the Figure. The ideal (infinite quantization levels, i.e. quantizer resolution, N = infinity) SNR for this condition is 81.02 dB full scale ("dBFS").

[0037] Figures 5(a) to (d) show the plot of output swing probability at the output of (a) first integrator, (b) second integrator, (c) quantizer, and (d) modulator of both circuits Figure 4(a) and Figure 4(b). It can be seen that with 0dB full-scale input signal, a conventional second-order modulator 400 which incorporates the quantizer output swing reduction method and circuit 410 of the present invention has a smaller quantizer output swing 501 compared to the one without 502. Figures 5(a)-(d) graphically illustrate that the present invention helps reduce the output swing at the output of second integrator 504. As seen therein, Figures 5(a)-(d) are plots of the probablity of occurrences at each voltage values of a sinewave input over a period of time. Figure 5(a) is a plot of the results from the first integrator output. Figure 5(b) is a plot of the results from the second integrator output. Figure 5(c) is a plot of the

results from the quantizer output. Figure 5(d) is a plot of the results from the modulator output of a second order sigma-delta modulator which implements the embodiment of the present invention seen in Figure 4(a). This plot also indicates the maximum output swing possible at these nodes in both circuits for a particular input signal amplitude/swing.

[0038] Figure 6 is a plot of the power spectrum that shows that the second-order modulator which incorporates the quantizer output swing reduction circuit 410 of the present invention has better SNR (80.9dBFS) performance, which is close to the ideal SNR (81.02dBFS) with 0dB full-scale input signal. The quantizer output swing of the conventional modulator 430 without the quantizer output swing reduction circuit 410 clips or is saturated before 0dB full-scale input signal due to the limited finite number of levels in the quantizer.

[0039] As seen in Figure 7, harmonic tones/distortion 701, tones/noise floor 702, and SFDR 703 is improved with the quantizer swing reduction circuit 410 of the present invention. The degradation of the SNR and SFDR performance due to quantizer output signal saturation in the second order sigma-delta modulator of Figure 4(b) compared to the one in Figure 4(a) which incorporates the present invention as a function of modulator input signal amplitude, are shown in Figure 7. The SNR and SFDR of the modulator 430 of Figure 4(b) without the quantizer swing reduction method and circuit of the present invention starts to drastically drop when the modulator input signal swings near 0dB full-scale compared to the modulator 400 with the quantizer output swing reduction method and circuit of the present invention.

[0040] Figure 8(a) illustrates using Yout as Vx to implement the proposed quantizer swing reduction method with a second-order sigma-delta modulator. The implementation of Figure 8(a) comprises first integrator 817 and a second integrator 803 with

transfer function $H1 = H2 = \frac{Z^{-1}}{1+Z^{-1}}$ and gain co-efficients, "a" 818 and "c" 821 equal to ½, and "b" 819 and "d" 820 equal to 2. The transfer function in the quantizer swing reduction block 810 is chosen to be identical to the one used in Figure 4(a) (seen as 410), such that $H3 = 1 - Z^{-1}$ and $H2 \cdot H3 = Z^{-1}$ in order to obtain the desired node voltage at the first integrator 817, output Y1 828, second integrator 803, output Y2 808, quantizer 805 output Y3 809, and modulator 800 output Yout 806, as expressed in equations [14], [15], [16] and [17] below:

$$Y1 = \frac{1}{2} \cdot \left[Vin \cdot Z^{-1} \cdot \left(1 + Z^{-1}\right) - \varepsilon_q \cdot Z^{-1} \cdot \left(1 - Z^{-1}\right) \right] \quad [14]$$

$$Y2 = Vin \cdot Z^{-2} \cdot \left(1 - Z^{-1}\right) + \varepsilon_q \cdot \left[\left(1 - Z^{-1}\right)^3 - 1 \right] \quad [15]$$

$$Y3 = Vin \cdot Z^{-2} \cdot \left(1 - Z^{-1}\right) + \varepsilon_q \cdot \left(1 - Z^{-1}\right)^3 \quad [16]$$

$$Yout = Vin \cdot Z^{-2} + \varepsilon_q \cdot \left(1 - Z^{-1}\right)^2 \quad [17]$$

Referring to equation [15], the signal reduction at quantizer 805 output Y3 809, is achieved by means of first order shaping of the original quantizer swing (i.e. both the two-delayed modulator 800 input 806 signal, Vin, term and the second order quantization noise, εq, shaping term).

[0041] The identical quantizer swing reduction circuit 410 of Figure 4(a) can also be applied to a first order sigma-delta modulator as seen in Figure 9(a). The node voltage at the integrator 903 Y2 output 908, the quantizer 905 output Y3 909, and modulator 900 output Yout 906, as expressed in equations [18] , [19] and [20] below:

$$Y2 = Vin \cdot Z^{-1} \cdot \left(1 - Z^{-1}\right) + \varepsilon_q \cdot \left[\left(1 - Z^{-1}\right)^2 - 1 \right] \quad [18]$$

$$Y3 = Vin \cdot Z^{-1} \cdot \left(1 - Z^{-1}\right) + \varepsilon_q \cdot \left(1 - Z^{-1}\right)^2 \quad [19]$$

$$Yout = Vin \cdot Z^{-1} + \varepsilon_q \cdot \left(1 - Z^{-1}\right) \qquad [20]$$

Referring to equation [18], the signal reduction at quantizer 905 output Y3 909 is achieved by means of first order shaping of original quantizer swing (i.e. both the delayed modulator 900 input 906 signal, Vin, term and first order quantization noise, εq, shaping term).

[0042] As can be seen, a primary advantage of the quantizer output swing reduction method and circuit of the present invention is that it enables the modulator to have a larger input/output swing range without degrading the SNR and SFDR performance due to quantizer output signal clipping or saturation.

[0043] The embodiments shown and described above are only exemplary. Even though numerous characteristics and advantages of the present invention have been set forth in the foregoing description together with details of the method of the invention, the disclosure is illustrative only and changes may be made within the principles of the invention to the full extent indicated by the broad general meaning of the terms used in the attached claims.